Discrete Applied Mathematics 156 (2008) 3216–3222

Contents lists available at ScienceDirect

Discrete Applied Mathematics

journal homepage: www.elsevier.com/locate/dam

A clique-difference encoding scheme for labelled k-path graphs

P.R.C. Pereira [a], L. Markenzon [b,*], O. Vernet [b]

[a] Instituto Militar de Engenharia, Praça General Tibúrcio, 80, 22290-270, Rio de Janeiro, Brazil

[b] Núcleo de Computação Eletrônica, Universidade Federal do Rio de Janeiro, CP 2324, 20010-974, Rio de Janeiro, Brazil

ARTICLE INFO

Article history:
Received 11 October 2005
Received in revised form 22 August 2007
Accepted 21 May 2008
Available online 9 July 2008

Keywords:
k-tree
k-path graph
Clique
Codeword

ABSTRACT

We present in this paper a codeword for labelled k-path graphs. Structural properties of this codeword are investigated, leading to the solution of two important problems: determining the exact number of labelled k-path graphs with n vertices and locating a hamiltonian path in a given k-path graph in time $O(n)$. The corresponding encoding scheme is also presented, providing linear-time algorithms for encoding and decoding.

1. Introduction

The idea of associating codewords with the labelled graphs of a specific family is not recent: in [1], Prüfer's article from 1918, a one-to-one correspondence between the set of $(n-2)$-tuples of the integers $\{1, 2, \ldots, n\}$ and the set of all labelled 1-trees on n vertices was already proved to hold. Rényi and Rényi [2] extended Prüfer's original codification to a broader family: the labelled rooted k-trees. In 1993, Chen [3] proposed a more compact code for k-trees, based on an intermediate representation using doubly labelled trees. More recently, Deo and Micikevicius [4] presented a survey on Prüfer-like codes for labelled trees, gathering and classifying several related codification algorithms. An alternative code for labelled k-trees was presented in [5], which has the same size as Chen's code but is much simpler to obtain.

A *codeword* for a certain graph is ultimately an alternative representation for this graph, usually more compact than the traditional edge-based ones, obtained by applying a set of *encoding rules*. The converse process, that is, constructing the graph corresponding to a given codeword is called *decoding*. An *encoding scheme* consists henceforth of encoding and decoding algorithms. By analyzing the properties of the codewords of all graphs belonging to a family under a given encoding scheme, it is sometimes possible to solve combinatorial problems for the family, such as counting, enumerating and randomly generating its members. Even algorithmic problems can be more efficiently solved when a graph is represented by a codeword.

Beineke and Pippert [6] introduced k-paths, generalizing the classic concept of simple paths. In [7], k-path graphs were defined and characterized: the family extends path graphs in the same way that k-trees do it for ordinary trees. In this paper we show that every k-path graph with n vertices can be assigned a sequence of $n-k-1$ pairs of vertices, named the *reduced sequence* and we investigate its structural properties. By presenting linear-time algorithms which allow unambiguous encoding and decoding, we show that the reduced sequence, along with the set of vertices, constitutes a codeword for a k-path graph that employs just $O(n)$ elements, where n is the number of vertices of the graph. Necessary and

* Corresponding author. Nucleo de Computacao Eletronica, Universidade Federal do Rio de Janeiro, CP.2324, 20010-974, Rio de Janeiro, Brazil. Tel: +55 21 25983256.
E-mail addresses: prenato@ime.eb.br (P.R.C. Pereira), markenzon@nce.ufrj.br (L. Markenzon), oswaldo@nce.ufrj.br (O. Vernet).

0166-218X/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.dam.2008.05.015

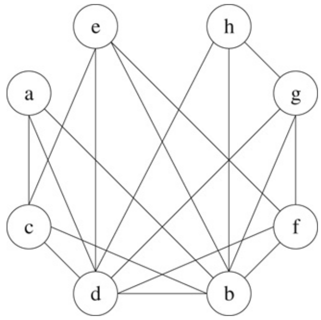

Fig. 1. A 3-path graph.

sufficient conditions for a sequence of pairs of vertices to be the reduced sequence of a k-path graph are obtained, yielding a validation theorem for such sequences. We also prove that k-path graphs are interval graphs and explore some properties in this context. Finally, two relevant problems are solved for the family: determining the exact number of k-path graphs with n vertices and locating hamiltonian paths in such graphs in $O(n)$-time.

2. *K*-paths and *k*-path graphs

In this section we summarize some definitions and results about k-paths, k-trees and k-path graphs. Basic results about chordal graphs and k-trees are assumed to be known and can be found in Blair and Peyton [8], Brandstädt [9] and Diestel [10].

Definition 1. A k-tree, $k > 0$, can be inductively defined as follows:

- Every complete graph with k vertices is a k-tree.
- If $G = (V, E)$ is a k-tree, $v \notin V$ and $Q \subseteq V$ is a k-clique of G, then $G' = (V \cup \{v\}, E \cup \{\{v, w\} | w \in Q\})$ is also a k-tree.
- Nothing else is a k-tree.

A simple path of length $p \geq 0$ in a graph G is an alternating sequence of distinct vertices and edges, $\langle v_0, e_1, v_1, e_2, \ldots, v_{p-1}, e_p, v_p\rangle$, so that $e_i = \{v_{i-1}, v_i\}$, $1 \leq i \leq p$. A graph consisting of a simple path on n vertices is called a *path graph* and is denoted as P_n. The concept of a k-path appeared first in [6], as a generalization of paths.

Definition 2. In a graph $G = (V, E)$, a k-path of length $p > 0$ is a sequence $\langle B_0, C_1, B_1, C_2, B_2, \ldots, C_p, B_p\rangle$, where:

- $B_i \subset V$, $0 \leq i \leq p$, are distinct k-cliques of G;
- $C_i \subseteq V$, $1 \leq i \leq p$, are distinct $(k + 1)$-cliques of G;
- $B_{i-1} \subset C_i$, $B_i \subset C_i$ and no other k-clique B_j, $0 \leq j \leq p$, $j \neq i - 1$ and $j \neq i$, is a subset of C_i, $1 \leq i \leq p$.

K-path graphs were first introduced in [7], along with a characterization theorem concerning the existence and exact number of simplicial vertices.

Definition 3 (*[7]*). Let $G = (V, E)$ be a k-tree with $n > k$ vertices. G is a k-path graph if there is a maximal k-path $\langle B_0, C_1, B_1, \ldots, C_p, B_p\rangle$, $p > 0$, such that the subgraph of G induced by $C_1 \cup \cdots \cup C_p$ is isomorphic to G.

Theorem 4 ([*7*]). *Let $G = (V, E)$ be a k-tree with $n > k+1$ vertices. G is a k-path graph if and only if G has exactly two simplicial vertices.*

Based on Theorem 4, the recognition of a k-tree $G = (V, E)$ as a k-path graph is straightforward: either it is a complete graph with $k + 1$ vertices or it has exactly two vertices with degree k, which are simplicial in G. In both cases, it suffices to examine $|\mathrm{Adj}(v)|$, $\forall v \in V$. The time complexity of this procedure is $O(m) = O(k\,n)$.

In a k-path graph G with n vertices, there are exactly k^2 maximal k-paths with length $n - k$. All these paths have the same subsequence of $(k + 1)$-cliques $\mathcal{C}(G) = \langle C_1, C_2, \ldots, C_{n-k}\rangle$, which is called the *core sequence* of G.

In Fig. 1, a 3-path graph with 8 vertices is depicted, whose core sequence is $\langle\{a, b, c, d\}, \{b, c, d, e\}, \{b, d, e, f\}, \{b, d, f, g\}, \{b, d, g, h\}\rangle$. Vertices a and h are simplicial.

Definition 5. A k-path graph can be inductively defined as follows:

- Every complete graph with $k + 1$ vertices is a k-path graph.
- If $G = (V, E)$ is a k-path graph, $Q \subset V$ is a k-clique of G containing at least one simplicial vertex and $v \notin V$, then the augmented graph $G' = (V \cup \{v\}, E \cup \{\{v, w\} \mid w \in Q\})$ is also a k-path graph.
- Nothing else is a k-path graph.

Recall that the *clique-intersection graph* of a graph G is the connected weighted graph whose vertices are the maximal cliques of G and whose edges connect vertices corresponding to non-disjoint maximal cliques. Each edge is assigned an integer weight, given by the cardinality of the intersection between the maximal cliques represented by its endpoints. Every maximum-weight spanning tree of the clique-intersection graph of G is called a *clique-tree* of G. If G is chordal, the clique-trees of G obey the *intersection property*: if Q_1 and Q_2 are maximal cliques of G, the intersection $Q_1 \cap Q_2$ is a subset of any maximal clique of G lying on the path between Q_1 and Q_2 in any clique-tree of G. See [8] for more details.

Lemma 6 shows an important property concerning the clique-tree of a k-path graph, which enables us to conclude that every k-path graph is an interval graph (see [11] for more details on interval graphs).

Lemma 6. *Every k-path graph admits a unique clique-tree, that is a path.*

Proof. From Definition 2, it is clear that, for $1 \leq i < p$, $|C_i \cap C_{i+1}| = k$, which is the maximum cardinality of the intersection between two maximal cliques of a k-path graph, since all of them have size $k+1$. Thus $(\{C_1, \ldots, C_p\}, \{\{C_i, C_{i+1}\} \mid 1 \leq i < p\})$ is a clique-tree of G, whose edges have weight k. As non-consecutive cliques have less than k vertices in common, this clique-tree is unique. □

By Lemma 6, the unique clique-tree of a k-path graph is a path whose nodes $C_1, \ldots, C_p$ appear in the same order as they do in the core sequence. By applying the intersection property on this unique clique-tree, we obtain the following lemma.

Lemma 7. *Let $\mathcal{C}(G) = \langle C_1, \ldots, C_p\rangle$ be the core sequence of a k-path graph G.*

- *If $v \in C_i - C_{i+1}$, then $v \notin C_j$, $1 \leq i < j \leq p$.*
- *If $v \in C_i - C_{i-1}$, then $v \notin C_j$, $1 \leq j < i \leq p$.*

Proof. If $v \in C_i$ and $v \in C_j$, where $1 \leq i < j \leq p$, then, according to the intersection property, v must also belong to $C_{i+1}, \ldots, C_{j-1}$, which contradicts $v \in C_i - C_{i+1}$. The second part can be proved analogously. □

3. The reduced sequence and its properties

In this section, we present the *reduced sequence* for k-path graphs, which is based on the difference between consecutive cliques of the core sequence. The theoretic framework introduced is fundamental to the development of the algorithms in the next section. The results lead to Theorem 14, which shows necessary and sufficient conditions for a given sequence of pairs to be the reduced sequence of a k-path graph.

Let $G = (V, E)$ be a k-path graph. Since there are k^2 maximal k-paths containing all $p = n - k$ cliques of size $(k + 1)$ of G, G cannot be univocally represented by one of these k-paths. However, given the core sequence $\mathcal{C}(G) = \langle C_1, C_2, \ldots, C_p\rangle$, it is possible to construct the corresponding graph unambiguously. Also the reverse sequence $\langle C_p, \ldots, C_2, C_1\rangle$ is an equivalent representation for the same k-path graph.

The $(k+1)$-cliques belonging to the core sequence have a peculiar structural property, proved in Lemma 8, which supports the definition of the reduced sequence.

Lemma 8. *If $\mathcal{C}(G) = \langle C_1, C_2, \ldots, C_p\rangle$, $p > 1$, is the core sequence of a k-path graph G with $n > k + 1$ vertices, then the sets $C_i - C_{i+1}$ and $C_{i+1} - C_i$, $1 \leq i < p$, are unitary.*

Proof. Since $|C_i| = |C_{i+1}| = k + 1$, $|C_i \cap C_{i+1}| = k$ and $C_i \neq C_{i+1}$. □

Based on the previous lemma, we can state the following definition.

Definition 9. Let $C_i - C_{i+1} = \{\ell_i\}$ and $C_{i+1} - C_i = \{r_i\}$, $1 \leq i < p$. The sequence $\mathcal{R}(G) = \langle(\ell_1, r_1), (\ell_2, r_2), \ldots, (\ell_{p-1}, r_{p-1})\rangle$ is called the *reduced sequence* of the k-path graph G. When $p = n - k = 1$, $\mathcal{R}(G) = \langle\,\rangle$ is empty.

As the reduced sequence $\mathcal{R}(G)$ is defined after the core sequence $\mathcal{C}(G)$, the mirrored sequence $\langle(r_{p-1}, \ell_{p-1}), \ldots, (r_1, \ell_1)\rangle$ also represents the same k-path graph G.

By Definition 9, the reduced sequence of the 3-path graph shown in Fig. 1 is $\langle(a, e), (c, f), (e, g), (f, h)\rangle$. It can be observed that not every vertex appears in the reduced sequence of a k-path graph (e.g. vertices b and d). These absent vertices belong to every $(k+1)$-clique of the graph, having degree $n-1$ and are so called *universal* vertices. By removing one of such vertices, we obtain a $(k-1)$-path graph with $n-1$ vertices, which has the same reduced sequence as the original graph. Fig. 2 shows this process applied to the k-path graph of Fig. 1. For this reason, a codeword based on the reduced sequence must also include the set of vertices, in order to allow the reconstruction of the corresponding k-path graph without ambiguities.

Given the reduced sequence of a k-path graph, its two simplicial vertices can be straightforwardly identified, as stated in Lemma 10.

Lemma 10. *Let $\mathcal{R}(G)$ be the reduced sequence of a k-path graph G. Then ℓ_1 and r_{p-1} are the simplicial vertices of G.*

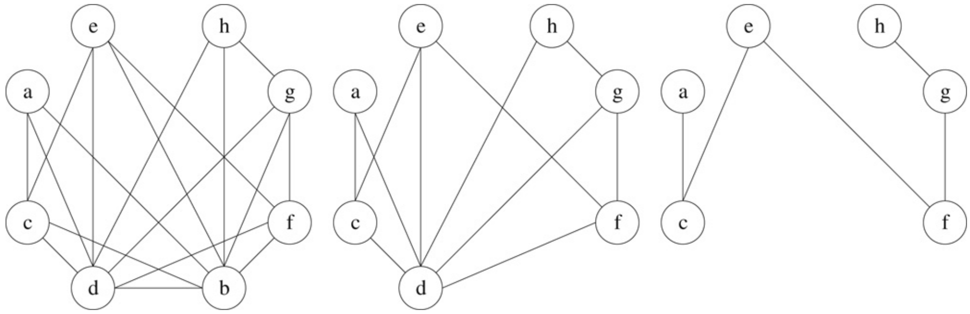

Fig. 2. k-path graphs with the same reduced sequence, $k = 3, 2, 1$.

Proof. Since $C_1 - C_2 = \{\ell_1\}$, $\ell_1 \in C_1$ and $\ell_1 \notin C_2$. Thus, by Lemma 7, ℓ_1 cannot belong to any other clique C_i, $1 < i \le p$, otherwise it should also belong to C_2. So $\mathrm{Adj}(\ell_1) = C_1 - \{\ell_1\}$, which is a k-clique. Likewise we can prove that r_{p-1} belongs only to C_p and that $\mathrm{Adj}(r_{p-1}) = C_p - \{r_{p-1}\}$. □

The removal of a simplicial vertex from a k-path graph originates a new k-path graph, whose core and reduced sequences are related to those of the original k-path graph as stated in Lemma 11.

Lemma 11. *Let $G = (V, E)$ be a k-path graph, $\mathcal{C}(G)$ its core sequence and $\mathcal{R}(G)$ its reduced sequence. Then*

- $\mathcal{C}(G - \ell_1) = \langle C_2, \ldots, C_p\rangle$ *and* $\mathcal{R}(G - \ell_1) = \langle(\ell_2, r_2), \ldots, (\ell_{p-1}, r_{p-1})\rangle$;
- $\mathcal{C}(G - r_{p-1}) = \langle C_1, \ldots, C_{p-1}\rangle$ *and* $\mathcal{R}(G - r_{p-1}) = \langle(\ell_1, r_1), \ldots, (\ell_{p-2}, r_{p-2})\rangle$.

Proof. Using Definition 5 and the previous lemma. □

The following lemma imposes constraints on the vertices appearing in the reduced sequence, that will be fundamental to the proof of Theorem 14.

Lemma 12. *Let $\mathcal{R}(G) = \langle(\ell_1, r_1), (\ell_2, r_2), \ldots, (\ell_{p-1}, r_{p-1})\rangle$ be the reduced sequence of a k-path graph G. Then $\ell_1 \ne \ell_2 \ne \cdots \ne \ell_{p-1}$, $r_1 \ne r_2 \ne \cdots \ne r_{p-1}$, $\ell_i \ne r_{j-1}$, for $1 \le i < j \le p$, and $\ell_i \ne r_{i-1}$, for $1 < i < p$.*

Proof. Let $1 \le i < j < p$. Since $\ell_i \in C_i$ and $\ell_i \notin C_{i+1}$, by Lemma 7, $\ell_i \notin C_j$. But $\ell_j \in C_j$. Hence, $\ell_i \ne \ell_j$. Similarly we can prove that $r_i \ne r_j$.

Now, let $1 \le i < j \le p$. Since $\ell_i \in C_i$ and $\ell_i \notin C_{i+1}$, $\ell_i \notin C_j$. But $r_{j-1} \in C_j$. Hence, $\ell_i \ne r_{j-1}$.

Finally, for $1 < i < p$, suppose that $\ell_i = r_{i-1}$. Then $C_i - C_{i+1} = C_i - C_{i-1}$. Since $|C_{i-1}| = |C_i| = |C_{i+1}| = k + 1$, we would have $C_{i-1} = C_{i+1}$, which contradicts Definition 2. □

Now it is possible to relate consecutive cliques in the core sequence to the pairs of the reduced sequence. The result, given by Lemma 13, will be crucial in the decoding process.

Lemma 13. *Let $\mathcal{C}(G) = \langle C_1, C_2, \ldots, C_p\rangle$ and $\mathcal{R}(G) = \langle(\ell_1, r_1), (\ell_2, r_2), \ldots, (\ell_{p-1}, r_{p-1})\rangle$, $p > 1$, be the core and the reduced sequences of a k-path graph $G = (V, E)$, respectively. Then $C_1 = V - \{r_1, \ldots, r_{p-1}\}$, $C_p = V - \{\ell_1, \ldots, \ell_{p-1}\}$ and $C_{i+1} = \{r_i\} \cup C_i - \{\ell_i\}$, for $1 \le i < p$.*

Proof. Since $\{r_i\} = C_{i+1} - C_i$, by Lemma 7, $r_i \notin C_1$, for $1 \le i < p$. Thus, if $v \in C_1$, then $v \ne r_i$ and $v \in V - \{r_1, \ldots, r_{p-1}\}$. Hence $C_1 \subseteq V - \{r_1, \ldots, r_{p-1}\}$. As $|C_1| = |V - \{r_1, \ldots, r_{p-1}\}| = k + 1$, the equality $C_1 = V - \{r_1, \ldots, r_{p-1}\}$ holds. Likewise we can prove that $C_p = V - \{\ell_1, \ldots, \ell_{p-1}\}$. Since $C_i - C_{i+1} = \{\ell_i\}$, $C_{i+1} - C_i = \{r_i\}$, $\ell_i \ne r_i$ and $|C_i| = |C_{i+1}| = k + 1$, we must have $C_{i+1} = \{r_i\} \cup C_i - \{\ell_i\}$. □

Theorem 14 states necessary and sufficient conditions for a sequence of $p = n - k$ pairs of vertices to be the reduced sequence of a k-path graph with n vertices.

Theorem 14. *Given $n, k, p = n - k$ and $S = \langle(\ell_1, r_1), (\ell_2, r_2), \ldots, (\ell_{p-1}, r_{p-1})\rangle$, $p > 2$, let $L_i = \bigcup_{j=1}^{i}\{\ell_j\}$ and $R_i = \bigcup_{j=i}^{p-1}\{r_j\}$. S is the reduced sequence of a k-path graph with n vertices if, and only if, $|L_i \cup R_{i-1}| = p + 1$, for $1 < i < p$.*

Proof. ($\Rightarrow$) By Lemma 12, the elements $\ell_1, \ldots, \ell_i, r_{i-1}, r_i, \ldots, r_{p-1}$ are all distinct, for $1 < i < p$. So $|L_i \cup R_{i-1}| = |L_i| + |R_{i-1}| = i + p - i + 1 = p + 1$, for $1 < i < p$.

($\Leftarrow$) Let V be any set such that $|V| = n$ and $\ell_i, r_i \in V$, for $1 \le i < p$. The proof uses induction on $|S|$:

· For $S = \langle(\ell_1, r_1), (\ell_2, r_2)\rangle$, $|L_2 \cup R_1| = 4$ and G consists of three maximal $(k + 1)$-cliques $C_1 = V - \{r_1, r_2\}$, $C_2 = V - \{\ell_1, r_2\}$ and $C_3 = V - \{\ell_1, \ell_2\}$.

· Suppose that the theorem is valid when $|S| = p - 2$.

· Consider $S = \langle(\ell_1, r_1), (\ell_2, r_2), \ldots, (\ell_{p-1}, r_{p-1})\rangle$ satisfying the assumptions. Thus $S' = \langle(\ell_1, r_1), (\ell_2, r_2), \ldots, (\ell_{p-2}, r_{p-2})\rangle$ satisfies $|L_i \cup R_{i-1}| = p + 1$, for $1 < i < p - 1$ and, by the inductive hypothesis, there exists a k-path graph $G' = (V', E')$, where $V' = V - \{r_{p-1}\}$, such that $\mathcal{R}(G') = S'$. Let $\mathcal{C}(G') = \langle C_1, \ldots, C_{p-1}\rangle$.
By Lemma 13, $C_{p-1} = V' - \{\ell_1, \ldots, \ell_{p-2}\}$. As $\ell_{p-1} \in V'$, we must have $\ell_{p-1} \in C_{p-1}$. By Lemma 10, r_{p-2} is simplicial in G' and also belongs to C_{p-1}.
Since $|L_{p-1} \cup R_{p-2}| = p + 1$, the elements $\ell_1, \ldots, \ell_{p-1}, r_{p-2}, r_{p-1}$ are all distinct. Let $G = (V, E' \cup \{\{r_{p-1}, v\} \mid v \in C_{p-1} - \{\ell_{p-1}\}\})$. Evidently G is a k-tree, where r_{p-1} is simplicial. However, r_{p-2}, which is simplicial in G', is not simplicial in G, because its two neighbors ℓ_{p-1} and r_{p-1} are not adjacent. ℓ_1 is simplicial in G' and remains simplicial in G, since $\ell_1 \notin C_{p-1}$. Hence, G has the same number of simplicial vertices as G': exactly 2. Thus, G is also a k-path graph, whose core sequence is $\langle C_1, \ldots, C_{p-1}, \{r_{p-1}\} \cup C_{p-1} - \{\ell_{p-1}\}\rangle$ and whose reduced sequence is S. □

4. The encoding and decoding algorithms

Due to the existence of universal vertices, the codeword for a k-path graph must consist of the set of vertices along with the reduced sequence. This section presents linear-time algorithms for obtaining the reduced sequence for a given k-path graph (encoding) and, conversely, reconstructing the graph from its codeword (decoding). The efficiency of these procedures is fundamental to endorse the usage of this codeword in algorithms involving k-path graphs.

Let us first consider the problem of coding a k-path graph, that is, given a k-path graph $G = (V, E)$ with $|V| = n$ vertices, we are interested in obtaining its reduced sequence $\mathcal{R}(G)$.

If $n = k + 1$, G is actually a complete graph and $\mathcal{R}(G) = \langle\rangle$. Otherwise G has two simplicial vertices, say s and t. In this case, the algorithm computes initially the degrees $d(v)$, for all $v \in V$, and identifies the simplicial vertices s and t as those having $d(s) = d(t) = k$. By Lemma 10, s and t correspond to the elements ℓ_1 and r_{p-1} of the reduced sequence, respectively. The remaining elements are determined by traversing G twice:

- The first traversal begins at s and removes it from G. By Lemma 11, the resultant graph $G - s$ is a k-path graph and has two vertices of degree k, which are simplicial: one of them is s', adjacent to s, and the other one is the very vertex t. By Lemma 11, s' corresponds to ℓ_2 in the reduced sequence. By repeating this step while the graph being analyzed has more than $k + 1$ vertices, the elements $\ell_2, \ldots, \ell_{p-1}$ are determined.
- The second traversal begins at t and is performed analogously, obtaining the elements $r_{p-2}, \ldots, r_1$.

The removals need not be explicitly performed; instead, the degrees of the vertices can be appropriately updated in order to emulate the removal process. Hence the time complexity of the coding algorithm is $O(|E|) = O(n\,k)$.

The problem of decoding can be stated as follows: given the set of vertices V and the reduced sequence $\mathcal{R}(G) = \langle(\ell_1, r_1), (\ell_2, r_2), \ldots, (\ell_{p-1}, r_{p-1})\rangle$ of a k-path graph G, construct E, the set of edges of G.

According to Lemma 13, the core sequence $\mathcal{C}(G) = \langle C_1, C_2, \ldots, C_p\rangle$ can be obtained by assigning $C_1 = V - \{r_1, \ldots, r_{p-1}\}$ and $C_{i+1} = \{r_i\} \cup C_i - \{\ell_i\}$, $1 \le i < p$. This leads to an algorithm for determining E:

- Initialize E with all edges whose extremities lie in $C = V - \{r_1, \ldots, r_{p-1}\}$;
- For $i = 1, \ldots, p - 1$, perform the following actions: $C = C - \{\ell_i\}$, $E = E \cup \{\{r_i, x\}\}, \forall x \in C$, and $C = C \cup \{r_i\}$.

Thus, obtaining the original graph from its codeword can also be done in time $O(|E|) = O(n\,k)$.

5. *K*-path graphs as interval graphs

Let $G = (V, E)$ be a k-path graph, $\mathcal{C}(G) = \langle C_1, \ldots, C_p\rangle$ its core sequence and $\mathcal{R}(G) = \langle(\ell_1, r_1), (\ell_2, r_2), \ldots, (\ell_{p-1}, r_{p-1})\rangle$ its reduced sequence. Since the unique clique-tree of G is a path, G is an interval graph and every vertex belongs to a range of consecutive cliques in $\mathcal{C}(G)$. Thus we can define the *scope* of each vertex v as the interval $[\alpha(v), \omega(v)]$, where $\alpha(v) = \min\{i \mid v \in C_i\}$ and $\omega(v) = \max\{i \mid v \in C_i\}$.

Through the assignment of these labels to every vertex of the k-path graph, it is possible to solve more efficiently some intermediate tasks that occur very often in the development of algorithms: determining the degrees of all vertices in $O(n)$-time and testing for the existence of an edge between a pair of vertices in $O(1)$-time.

Initially, we show how the labels $\alpha(v)$ and $\omega(v)$ can be deduced from the reduced sequence. In Lemma 15, the relation between the occurrence of a vertex in the reduced sequence and its scope is explored.

Lemma 15. *Let $v \in V$. The following mutually exclusive situations arise:*

- $\alpha(v) = 1$ *and* $\omega(v) = p$ *if and only if* $v \ne \ell_i$ *and* $v \ne r_i$, *for* $1 \le i < p$;
- $\alpha(v) = 1$ *and* $\omega(v) < p$ *if and only if* $v = \ell_{\omega(v)}$ *and* $v \ne r_i$, *for* $1 \le i < p$;
- $\alpha(v) > 1$ *and* $\omega(v) = p$ *if and only if* $v \ne \ell_i$ *and* $v = r_{\alpha(v)-1}$, *for* $1 \le i < p$;
- $\alpha(v) > 1$ *and* $\omega(v) < p$ *if and only if* $v = \ell_{\omega(v)} = r_{\alpha(v)-1}$.

Proof. If $\alpha(v) = 1$, then $v \in C_1$. As, by Lemma 13, $C_1 = V - \{r_1, \ldots, r_{p-1}\}$, so $v \ne r_i$, for $1 \le i < p$. If $\alpha(v) > 1$, then $v \in C_{\alpha(v)}, \ldots, C_{\omega(v)}$ and $v \notin C_{\alpha(v)-1}$. Since $C_{\alpha(v)-1} - C_{\alpha(v)} = \{r_{\alpha(v)-1}\}$, then $v = r_{\alpha(v)-1}$. The proof for $\omega(v)$ follows analogously. □

As long as G is represented through its reduced sequence, Lemma 15 yields an $O(n)$-time algorithm for determining $\alpha(v)$ and $\omega(v)$ for every vertex of G.

The next lemma expresses the degree of a vertex v, $\delta(v)$, in terms of $\alpha(v)$ and $\omega(v)$.

Lemma 16. *Each vertex $v \in V$ satisfies $\delta(v) = \omega(v) - \alpha(v) + k$.*

Proof. The proof follows by induction on $|V|$.

- For $|V| = k+1$, the corresponding k-path graph is a complete graph with $k+1$ vertices. Thus $\delta(v) = k, \alpha(v) = \omega(v) = 1$ and the lemma holds.
- Suppose that the result is valid when $|V| = n$.
- Let $G = (V, E)$ with $|V| = n$ vertices and $t \notin V$. By the inductive assumption, $\delta(v) = \omega(v) - \alpha(v) + k$, for $v \in V$. Form $G' = (V \cup \{t\}, E')$, with $n + 1$ vertices, by joining t to a k-clique Q of G containing at least one simplicial vertex. We have to prove that, for $v \in V \cup \{t\}$, $\delta'(v) = \omega'(v) - \alpha'(v) + k$.
  - For $v \in V - Q$, $\alpha'(v) = \alpha(v)$ and $\omega'(v) = \omega(v)$, since t is not a neighbor of v. So the result holds.
  - For $v \in Q$, the degree of v is augmented by 1 due to the addition of t. Since $\{t\} \cup Q$ establishes a new maximal $(k + 1)$-clique in G', $\omega'(v) = \omega(v) + 1$ and $\alpha'(v) = \alpha(v)$. So $\delta'(v) = \delta(v) + 1 = \omega(v) - \alpha(v) + k + 1 = \omega'(v) - \alpha'(v) + k$.
  - Finally, as $\alpha'(t) = \omega'(t)$, $\delta'(t) = \omega'(t) - \alpha'(t) + k = k$. □

Thus, by the relation established in Lemma 16, the degrees of all vertices can also be computed in $O(n)$-time.

Testing for the existence of an edge between a given pair of vertices is a supporting operation in many algorithms, that can be accomplished in $O(1)$-time when the graph is represented by its adjacency matrix, requiring $O(n^2)$ space. However, for a k-path graph, having determined α and ω for each vertex, this condition can be tested in $O(1)$-time, avoiding the adjacency matrix, according to Lemma 17.

Lemma 17. *Let $v, w \in V$. $\{v, w\}$ is an edge of a k-path graph $G = (V, E)$ if and only if $\alpha(v) \le \omega(w)$ and $\alpha(w) \le \omega(v)$.*

Proof. It suffices to notice that the interval that must be assigned to each vertex $v \in V$ in order to characterize G as an interval graph is exactly the scope of v: $[\alpha(v), \omega(v)]$. Thus, $\{v, w\} \in E$ when the scopes of v and w are non-disjoint intervals. □

6. Counting labelled k-path graphs

An important application of the reduced sequence is to obtain the total number of labelled k-path graphs with n vertices, as stated in Theorem 18.

Theorem 18. *There are exactly*

$$\frac{n!}{k!} \times \frac{k^{n-k-2}}{2}$$

labelled k-path graphs with n vertices.

Proof. We begin with an empty reduced sequence and make successive choices of vertices in order to fill it up. The condition $|L_i \cup R_{i-1}| = p+1$, for $1 < i < p$, established in Theorem 14, helps us to guarantee that, at every step, the sequence represents a k-path graph.

Initially a subsequence of $p + 1 = n - k + 1$ distinct elements must be selected from among the n available vertices, yielding $\ell_1, \ell_2, r_1, \dots, r_{p-1}$. There are $\frac{n!}{(k-1)!}$ ways of performing this choice.

The remaining elements $\ell_3, \dots, \ell_{p-1}$ are successively chosen. When choosing ℓ_i, the elements $\ell_1, \dots, \ell_{i-1}, r_1, \dots, r_{p-1}$ have already been selected. Thus, there are exactly $n - p = k$ possibilities of choice for each ℓ_i. These choices occur $p - 3 = n - k - 3$ times, so we have k^{n-k-3} possibilities.

Since $\langle(\ell_1, r_1), \dots, (\ell_{p-1}, r_{p-1})\rangle$ and $\langle(r_{p-1}, \ell_{p-1}), \dots, (r_1, \ell_1)\rangle$ represent the same k-path graph, the total number of k-path graphs with n vertices is given by

$$\frac{n!}{(k-1)!} \times \frac{k^{n-k-3}}{2} = \frac{n!}{k!} \times \frac{k^{n-k-2}}{2}. \quad \square$$

7. Hamiltonian paths in k-path graphs

In this last section, the codeword defined for k-path graphs acts as an alternative data structure for representing such graphs and provides a straightforward solution to a very important graph-theoretical algorithmic problem: locating hamiltonian paths, according to Theorem 19.

Theorem 19. *Let $G = (V, E)$ be a k-path graph with n vertices and $\mathcal{R}(G) = \langle(\ell_1, r_1), (\ell_2, r_2), \ldots, (\ell_{p-1}, r_{p-1})\rangle$ its reduced sequence. Then G has $k!$ hamiltonian paths starting at ℓ_1 and containing the subsequence $\langle r_1, r_2, \ldots, r_{p-1}\rangle$ as a suffix.*

Proof. Let us prove first that the sequence $\langle r_1, r_2, \ldots, r_{p-1}\rangle$ is a simple path in G. By Lemma 12, $r_1 \neq r_2 \neq \cdots \neq r_{p-1}$. To see that $\{r_i, r_{i+1}\} \in E$, remember that $\{r_{i+1}\} = C_{i+2} - C_{i+1}$. So $r_{i+1} \in C_{i+2}$. Since $C_{i+2} = \{r_{i+1}\} \cup C_{i+1} - \{\ell_{i+1}\}$, $\{r_i\} = C_{i+1} - C_i$, $r_i \neq r_{i+1}$ and $r_i \neq \ell_{i+1}$, r_i must also belong to C_{i+2}. Thus there exists an edge joining r_i to r_{i+1}.

Moreover, by Theorem 14, $V - \{r_1, \ldots, r_{p-1}\} = C_1$ and the vertices belonging to $C_1 - \{\ell_1\}$ are neighbors of r_1. Thus there are $k!$ hamiltonian paths in G consisting of the concatenation of two simple paths: a prefix containing all vertices of C_1, starting at ℓ_1, and the suffix $\langle r_1, r_2, \ldots, r_{p-1}\rangle$. $\square$

Hence, if a k-path graph $G = (V, E)$ is represented by V and $\mathcal{R}(G)$, each of the aforementioned hamiltonian paths can be found in time $O(n)$.

Acknowledgement

The second author was partially supported by grant 306893/2006-1, CNPq, Brazil.

References

[1] A. Prüfer, Neuer Beweis eines Satzes über Permutationen, Archiv der Mathematik und Physik 27 (1918) 142–144.
[2] C. Rényi, A. Rényi, The Prüfer code for k-trees, in: P. Erdös, et al. (Eds.), Combinatorial Theory and its Applications, 1970, pp. 945–971.
[3] W.Y.C. Chen, A coding algorithm for Rényi trees, Journal of Combinatorial Theory Series A 63 (1993) 11–25.
[4] N. Deo, P. Micikevicius, Prüfer-like codes for labeled trees, Congressus Numerantium 151 (2001) 65–73.
[5] P.R.C. Pereira, L. Markenzon, O. Vernet, The reduced Prüfer code for rooted labelled k-trees, Electronic Notes in Discrete Mathematics 22 (2005) 135–139.
[6] L.W. Beineke, R.E. Pippert, Properties and characterizations of k-trees, Mathematika 18 (1971) 141–151.
[7] L. Markenzon, C. Justel, N. Paciornik, Subclasses of k-trees: Characterization and recognition, Discrete Applied Mathematics 154 (2006) 818–825.
[8] J.R.S. Blair, B. Peyton, An introduction to chordal graphs and clique trees, in: Graph Theory and Sparse Matrix Computation, IMA 56 (1993) 1–29.
[9] A. Brandstädt, V.B. Le, J.P. Spinrad, Graph classes — A survey, in: SIAM Monographs in Discrete Mathematics and Applications, 1999.
[10] R. Diestel, Graph Theory, 2nd edition, Springer Verlag, New York, 2000.
[11] M. Habib, R. McConnell, C. Paul, L. Viennot, Lex-BFS and partition refinement, with applications to transitive orientation, interval graph recognition and consecutive ones testing, Theoretical Computer Science 234 (2000) 59–84.